



Санкт-Петербургский  
Государственный  
Политехнический  
Университет

Институт прикладной  
математики и механики

**Научный дискусс**  
**«Может ли машина мыслить»**  
**Лекция 12**  
**Обучение - виртуальная основа**  
**«умных поступков»**

---

7 декабря  
2017 г.

## Содержание:

- Супервентность – синхронизация реального и виртуального
- Физическое vs информационное ....
- Что меняется в мозгу, когда мы «понимаем что-либо».

# Что обсуждалось на прошлой лекции



- Мозг – это система управления «химической машиной» организма человека.
- Эмоции ( мысли) производит в организме определенный химический фермент; Это и есть химический «путь» преобразования «мысли» в физическую реальность.
- Мозг как «машина управления химическими процессами » **меняется** в процессе своей деятельности, но на «химическом» уровне машина не отличает реальное, т.е. материальное, от виртуального, т.е. от воображаемого!
- **Различие формируется – на интеллектуальном уровне.**
- Где этот уровень находится и как он формируется ? Кто этим уровнем управляет ?

**Интеллект - Intellect-intelligence-mentality.** Ум, понятливость, разведка - a term used in studies of the **human mind**, and refers to the **ability of the mind** to come to correct conclusions about **what is true** or real, and about how to solve problems

**Разум – mind** (лат. ratio) - высший тип мыслительной деятельности ; по смыслу соответствует латинскому слову «**intellectus**» – понимание, способность анализа, абстрагирования и обобщения

**Сознание – consciousness** состояние человека, выражающееся в субъективном переживании **событий внешнего мира** и жизни самого индивида, а также в отчёте об этих событиях

**Мышление – thinking** - познавательная деятельность человека

# artificial intelligence - machine intelligence

В этом контексте **Intelligence** понимается как способность к «умным поступкам или делам» - smart deeds or doing

Основной вопрос: где этот «умный» уровень находится и как он формируется ?

**Ум - это индивидуальная особенность личности, выражающаяся в способности:**

- Мышления.
- Понимания.
- Думать.
- Находить решение задач.
- Предсказывать (предвидеть) последствия своих поступков.
- Использовать готовые решения.

**Ум состоит из:**

- Суммы усвоенных знаний.
- Интеллекта (умения использовать знания умело и с пользой).
- Интуиции (умения оперировать усвоенными знаниями на бессознательном уровне).
- Личностных установок, верований, убеждений, позиций, программ.

# Ум как «виртуальная реальность»



«Общие умственные способности зависят от сети внутри мозга, это связь серого вещества, или клеток мозга, с белым веществом, или соединительными волокнами между нейронами

*Proceedings of the National Academy of Sciences*

Гипотеза: «мозг состоит из элементарных частиц, которые существуют как игра воображения, они воображаемы».

Вопрос: «Ум существует в мозге или мозг воображаем умом?»... человек в целом воображаемая конструкция ума - > по теореме Геделя «система познать саму себя не может» .... Ум — это то, из чего мы созданы.

# Создание – синхронизации «реального и виртуального» через принятия рациональных решений

Сознание – это способность субъекта самостоятельно принимать **рациональные и моральные решения**

Понятие **Супервентности** (англ. Supervenience) – отношение детерминированности состояния любой системы состоянием другой системы.

Итак, набор свойств одной системы **супервентен** относительно набора свойств другой системы в том случае, если существование различия между двумя фактами в свойствах первой системы невозможно без существования такого же различия между двумя фактами в свойствах второй системы.

Понятие супервентности является центральным понятием современной **теории сознания** - описания объективной зависимости ментальных явлений от физических явлений,

«supervenience» – это «действие, возникающее как следствие чего-либо другого»

примеры:

- **отсутствие различий** в ментальных свойствах при отсутствии различий в физических свойствах;
- **отсутствие различий** в компьютерной программе при отсутствии различий в аппаратной конфигурации компьютера;
- **отсутствие различий** в экономике при отсутствии различий в поведении экономических агентов.

# Физическое vs информационное

**Тезис 1.** Если имеется два физически тождественных состояния системы, но в одном случае **мы** ничего не знаем о ее состоянии, а во втором имеем какую-то информацию об этом, то эти два состояния системы различаются **фундаментально**.

**Суть различия:** Во втором случае мы можем «заставить» систему совершать «работу», а в первом – нет! Принцип Ландауэра: в процессе потери информации о состоянии совершается работа.

**Информация материальна:** связана с физическими объектами (носителями)

**Материя информационна:** информация атрибут структуры материи, которая является носителем отличий для физических объектов – difference that make the difference.

**Тезис. 2 ( антропный принцип).** То что называется **объективные** характеристики физического мира связано с **существованием наблюдателя**, который воспринимает **различия между** объектами материального мира или информацию

# Знания субъекта и объективная реальность мира

Следствие из **Тезиса 2**: У открытой физической системы есть внешний по отношению к ней **наблюдатель** (познающий субъект), который может иметь или не иметь **знания** об объективных (реальных) свойствах этой системы.

По отношению к познающему субъекту все **проявления реальности** должны явно или неявно учитывать **объем его знаний** о системе.

Принципиальное ограничение объема **доступных для познающего субъекта знаний** (ресурсы памяти) о изучаемой физической системе определяет модельный характер законов, с помощью которых изучается ее поведения. Эти законы, в частности, называются:

- Термодинамика

- Теория относительности

- Квантовая механика

Эти законы, отражают ту точность наблюдения за природой, которую доставляют органы чувств или используемые средства измерений.

# Свойства информационной реальности – то, с чем «работает» интеллект субъекта познания

Все есть число  
Пифагор

Истина в неполноте  
Гедель

It from bit  
Арчибальд Уильер

В Природу можно понять, «рассмотрев» ее свойства либо с помощью органов чувств, либо приборов, (рентгеновских лучей ...) либо ... вычислить свойства с помощью математических моделей.

На прото-уровне материя обладает свойствами, которые непосредственно выражаются через числовые отношения, а эти выражения, в свою очередь, изоморфны программам вычислений.

**Тезис. 3.** Вычисления это физический процесс надо носителями информации – результат вычислений информационная реальность, существующая одновременно в физической и виртуальной (информационной) ипостасях.

# Проблема науки – кодирование знаний для «переноса» их во времени и пространстве.

Вплоть до XX века приборы, используемые физиками, давали информацию лишь о **экстенсиональных** или внешних признаках проявления отличий объектов природы. Экстенсивная мера информации **кодируется конечным набором цифр**, которые легко переносятся на:

- чернила на бумаге;
- электромагнитное поле;
- костяшки на счетах;
- положение стрелки прибора на шкале значений;

Однако объекты природы наделены также **интенсиональными** (скрытыми) характеристиками, **которые не могут непосредственно воздействовать на органы чувств**

- проницаемость мембран в нейронах;
- электрический потенциал;
- намагниченность;
- и т.д.

Можно ли и чем можно кодировать **скрытые** характеристики ?

# «difference that make the difference» ?!

**Тезис 4.** Информационная реальность это то, что «может обнаружить» себя в «пространстве-времени». Объекты такой реальности могут находиться как в «**ЧИСТЫХ**» состояниях, непосредственно доступных для измерения, так и в :

- В «**СМЕСИ**» состояний, т.е. **конкретном состоянии**, параметры которого зависят от параметров «чистых» состояний составных его частей (описываются т.н. матрицей плотности)
- В **суперпозиции** – линейной комбинации «чистых» состояний, т.е. находиться **одновременно во всех** «чистых» состояниях, но ... с различной вероятностью, результат измерения «коллапс» волновой функции дает случайный результат.
- **несепарабельной суперпозиции** «чистых» состояний, т.е. запутанной (взаимозависимой) суперпозиции, которая не факторизуется на тензорные произведения «чистых» состояний. В этих условиях возможен «обратный коллапс» волновой функции.

# Энтропия физической реальности

**«Физичность»** информации заключается в том, что она:

- отражает как свойства физической системы, которые доступны для непосредственного измерения с помощью приборов (**экстенсионал в форме температуры, проводимости...**), например

$$\Delta S_{ph} = \Delta Q/T \text{ [дж/К];}$$

- Так и выражается через число микросостояний системы (**интенсионал**) :

$$S_{inf} = \Omega (k \ln 2)$$

- В итоге обобщенная энтропия «физической реальности» имеет вид

$$S_{ph} + S_{inf}$$

где  $\Delta S_{ph} = \Delta Q/T$  [дж/К] – функция состояния равновесной системы, а статистическая величина, связывающая микроскопические состояния и термодинамические свойства, равна  $S_{inf} = \Omega (k \ln 2)$ , где  $\Omega$  - число **микросостояний** (способов), с помощью которых можно составить рассматриваемое **макроскопическое** состояние

# Принцип Ландауэра

**Принцип:** в процессе потери информации о состоянии системы совершается работа.

Потеря (стирание) одного бита информации сопровождается «потерей» свободной энергии по крайней мере  $kT \ln 2 = 0.693kT$  [Дж/К] ... ( $2.87 \cdot 10^{-21}$  Дж = 0.018 эВ при 300 К, где постоянная Больцмана  $k$  равна  $1,38 \cdot 10^{-23}$  Дж/К.)

Одновременно с потерей одного бита информации происходит увеличение энтропии окружающей среды на  $k \ln 2 = 0.96 \cdot 10^{-23}$  Дж/К. [  $S=Q/T$  ]

При этом в окружающую среду «передается» определенное количество теплоты

**Эти количества не могут быть уменьшены, какие бы способы записи, хранения или стирания информации ни применялись**

# Комментарии к принципу Ландауэра

- Замечание 1 (теоретическое). Принцип Ландауэра связывает информацию с классическими физическими концепциями энергии и энтропии.
- Замечание 2 (практическое). В настоящее время в реальности затраты энергии на элементарные вычислительные операции, осуществляемые с помощью электронных микропроцессоров намного больше ландауэровского предела ( $2.87 \cdot 10^{-21} \text{ Дж} = 0.018 \text{ эВ}$  при 300 К), принцип подсказывает возможные пути и пределы минимизации энергопотребления компьютеров
- Замечание 3 (Принцип Ландауэра есть следствие логической необратимости процессов потери информации. Аналогичные выводы справедливы и в отношении других необратимых логических операций, например, И, ИЛИ: **диссипация энергии составляет как минимум  $kT \ln 2$  на каждый потерянный бит**

1. Надо ли всегда и как исправлять ошибки...чтобы получить правильный результат вычислений ?
2. Можно ли создать «новую» математику, с помощью которой используя много ошибочных результатов можно получить практически верный результат ?
3. Как можно «автоматизировать» процесс сравнения между собой «сложных» объектов, используя
  - интенционал – смысл понятия, сопоставляемого объекту
  - экстенционал – объем понятия, сопоставляемого объекту